The Contact between Sinhala and English Orthography in Online Text Messages

Nimasha Malalasekera

Department of English Language Teaching, University of Kelaniya, Sri Lanka
nimashamalalasekera@gmail.com

Abstract— At present, using English characters to represent Sinhala words appears to be a common feature in online texting among Sinhala-English bilinguals. This study examines the contact between Sinhala and English orthography in online text messages in which Sinhala words are represented using English letters. Much research has been conducted on the contact between Sinhala and English, features and conventions of Internet language and language contact in digital Internet genres. However, no research has yet examined the contact between Sinhala and English orthography in digital Internet genres in an attempt to identify patterns arising in such linguistic phenomenon. The data for the study come from a corpus of online text messages sent via mobile messaging applications (hereafter apps) such as Facebook Messenger, WhatsApp and Viber collected from 20 undergraduates in the Faculty of Arts of the University of Colombo. The text messages were collected using the friend of a friend method. The study identifies four main patterns of representing Sinhala vowel sounds and two main patterns of representing Sinhala consonant sounds through English letters. The findings invoke a need to expand the conception of frequentional copying showing a potential emergence of a new variety of Sinhala in which the orthographic system is a result of the contact between Sinhala and English orthography. The identified patterns also help develop more user-friendly online Sinhala-English transliteration software than the existing ones such as Google Input Tools, SinGlish Transliterated and SinGlish (Phonetic) Transliterated which are either unable to identify certain patterns of representing Sinhala vowel and consonant sounds through English characters or have their own transliteration schemes which the users have to adopt. Thereby, these patterns challenge the conventional idea that no system of transliteration of a (more) phonetic language by an unphonetic one like English can be perfect.

Keywords: online texting, contact, English orthography, spoken Sinhala sounds

I. INTRODUCTION

Based on a corpus of online text messages\(^1\) sent via mobile messaging applications\(^2\) such as Facebook Messenger, WhatsApp and Viber in which Sinhala words are represented using English characters, this study aims to examine the results of the contact\(^3\) between Sinhala and English orthography in online text messages. It identifies patterns of representing spoken Sinhala vowel and consonant sounds\(^4\) through English letters\(^5\) in online texting since “electronic discourse is writing that very often reads as if it were being spoken” (Crystal, 2004, p. 25). Due to space constraints, the study will not focus on the identification of patterns of representing Sinhala vowel and consonant clusters through English characters.

Research has been conducted on “digital Internet genres” (Mendis, 2006, p. 125) focusing on features and conventions of Internet Language (See, for instance, Porter, 1996; Davis & Brewer, 1997; Baron, 2000; Crystal, 2004; Ziten & Stein, 2004 etc.) and the contact between two languages in digital Internet communication (See, for instance, Su, 2006; van Gass, 2008; Devic, 2008; Huang, 2009; Temur & Vuruş, 2009; Themistocleous, 2010; Tagg & Sargeant, 2012; Rafi, 2013 etc.). The contact between Sinhala and English has also been researched in the study of Ceylon English, Lankan English and/or Sri Lankan English (SLE) (See, for instance, Pasé, 1948, 1950, 1955; Halverson, 1966; Chithra Fernando, 1977; Siromi Fernando, n.d., 1985, 2003, 2008a, 2008b, 2011/2012; Gunesekara, 2000, 2008; Herat, 2006; Meyler, 2007; Siromi Fernando, Gunesekara & Parakrama, 2010 etc.). Moreover, based on Short Message Service (SMS) and print media, Widyalankara (2015) looks at the representation of Sinhala sounds through English characters referring to it as “interlingual texting.” Yet, her study claims that such representation leads to “ambiguity” and “unintelligibility” of the messages. No study has yet focused on the contact between the orthographic systems of Sinhala and English in digital Internet genres in an attempt to identify patterns emerging in such phenomenon, which leads to the research problem of the study:

How does Sinhala and English orthography come into contact in the use of English letters to represent Sinhala words in online text messages sent via mobile messaging apps such as Facebook Messenger, WhatsApp, and Viber?

1. What are the patterns of representing Sinhala vowel sounds in the word initial, medial and final positions in online text messages?

2. What are the patterns of representing Sinhala consonant sounds in the word initial, medial and final positions in online text messages?

The following are the objectives of the study:


2. To show the potential emergence of a new variety of Sinhala in which the orthographic system is a result of the contact between
Sinhala and English orthography in online texting.

3. To challenge Gunasékara’s (1962) claim about the systems of transliteration of a (more) phonetic language by unphonetic English.

4. To present an alternative viewpoint to Widyalankara’s (2015) claim about the “ambiguity” and “unintelligibility” of Sinhala messages represented using English characters.

5. To help develop more user-friendly Sinhala-English transliteration software.

II. METHODOLOGY AND EXPERIMENTAL DESIGN

The data include online text messages sent through mobile messaging apps namely, Facebook Messenger, WhatsApp and Viber. These were collected from 20 undergraduates in the Faculty of Arts of the University of Colombo over a period of 10 months. A version of Milroy’s (1980) “friend of a friend” method was used to collect data in order to protect the privacy of the authors of the messages. Two friends of the researcher who had more access to the sample than the researcher were requested to obtain messages from their friends. The sample was selected based on the convenient sampling technique presuming that many individuals would be unwilling to share their messages due to the level of privacy attached to them.


III. RESULTS

The patterns of representing Sinhala vowel sounds using English letters in the word initial, medial and/or final positions can be categorized as follows:6

A. The Representation of Spoken Sinhala Vowel Sounds

1. The use of (an) English symbol(s) to represent a Sinhala vowel sound in the word initial, medial and/or final positions where the particular symbol(s) is/are used to represent the same or a similar sound in Sri Lankan English. Below are some examples:

<table>
<thead>
<tr>
<th>English symbols used to represent the Sinhala vowel sound</th>
<th>Position in Sinhala words</th>
<th>Sinhala words</th>
<th>English words in which the said symbol produces the same or a similar sound</th>
<th>Position in English words</th>
</tr>
</thead>
<tbody>
<tr>
<td>a initial /æ:tin/ සේටින්</td>
<td>sand /sænd/</td>
<td>medial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kama /kæ:ma/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ba /bæ:/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. The use of (an) English symbol(s) to represent a Sinhala vowel sound in the word initial, medial and/or final positions where the particular English symbol(s) is/are used to represent (a) different sound(s) in Sri Lankan English. Below are some examples:

<table>
<thead>
<tr>
<th>English symbols used to represent the Sinhala vowel sound</th>
<th>Position in Sinhala words</th>
<th>Sinhala words</th>
<th>English words in which the said symbol produces the same or a similar sound</th>
<th>The sounds produced by the said symbols in English words</th>
</tr>
</thead>
<tbody>
<tr>
<td>ethin /æ:tin/ සේටින්</td>
<td>/e/- get</td>
<td></td>
<td></td>
<td>/e/- get</td>
</tr>
<tr>
<td>keli /kæ:li/ අල්කළි</td>
<td>/i/- kiss</td>
<td></td>
<td></td>
<td>/i/- kiss</td>
</tr>
<tr>
<td>be /bæ:/</td>
<td>/i:/ - mete</td>
<td></td>
<td></td>
<td>/i:/ - mete</td>
</tr>
<tr>
<td>aethin /æ:tin/ සේටින්</td>
<td>/ae/- jaeger</td>
<td></td>
<td></td>
<td>/ae/- jaeger</td>
</tr>
<tr>
<td>nae /næ:/</td>
<td></td>
<td></td>
<td></td>
<td>/dʒe:gə:/</td>
</tr>
</tbody>
</table>
3. The reduplication of an English symbol to represent a Sinhala vowel sound in the word initial, medial and/or final positions where the reduplication of the particular symbol represents a different sound in Sri Lankan English. Below are some examples:

### Table 3. Spoken Sinhala long back mid vowel 'ඉ' /o:/

<table>
<thead>
<tr>
<th>Reduplicated English symbol and spelling pattern</th>
<th>Sinhala words and positions in Sinhala words</th>
<th>English words in which the said symbols produce the spoken Sinhala vowel sound</th>
<th>The sound(s) produced by the said symbol(s) in English words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol: o</td>
<td>initial /o:lu/</td>
<td>/u/ good /gud/</td>
<td></td>
</tr>
<tr>
<td>Spelling pattern: oo</td>
<td>final /apo:/</td>
<td>/u:/ ooze /u:/z/ foo /fu:/ too /tu:/</td>
<td></td>
</tr>
</tbody>
</table>

4. The creation of a new spelling pattern that is not present in English to represent a Sinhala vowel sound in the word initial, medial and/or final positions through the reduplication of an English symbol. There are two types. Below are the two types with examples:

4.1. The use of the particular symbol twice.

### Table 4. Spoken Sinhala long front low vowel 'ඉ' /æ:/

<table>
<thead>
<tr>
<th>Reduplicated English symbol</th>
<th>New spelling pattern created</th>
<th>Position in Sinhala words</th>
<th>Sinhala words</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>aa</td>
<td>initial</td>
<td>aathin /æ:tin/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>median</td>
<td>kaalla /kæ:ll/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>final</td>
<td>baa /bæ:/</td>
</tr>
</tbody>
</table>

4.2. The use of the particular symbol more than twice.

### Table 5. Spoken Sinhala long front low vowel 'ඉ' /æ:/

<table>
<thead>
<tr>
<th>Reduplicated English symbol</th>
<th>New spelling pattern created</th>
<th>Sinhala words</th>
<th>Position in Sinhala words</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>aaa</td>
<td>/næ:/</td>
<td>final</td>
</tr>
</tbody>
</table>

B. The Representation of Spoken Sinhala Consonant Sounds

The patterns of representing Sinhala consonant sounds using English letters in the word initial, medial and/or final can be categorized as follows:

1. The use of (an) English symbol(s) to represent a Sinhala consonant sound in the word initial, medial and/or final positions where the particular symbol(s) produce(s) the same or a similar sound in Sri Lankan English. Below are some examples:

### Table 6. Spoken Sinhala dental unvoiced stop 'ත' /t/ |

<table>
<thead>
<tr>
<th>English symbols used</th>
<th>Position in Sinhala words</th>
<th>Sinhala words</th>
<th>English words in which the particular symbol(s) produce(s) the same or a similar sound</th>
<th>Position in English words</th>
</tr>
</thead>
<tbody>
<tr>
<td>th</td>
<td>initial</td>
<td>thibba /tʰiba:/</td>
<td>thin /tʰin/</td>
<td>initial</td>
</tr>
<tr>
<td></td>
<td>medial</td>
<td>aththa /aṱʰta:/</td>
<td>loathsome /lə:tsəm/</td>
<td>medial</td>
</tr>
<tr>
<td></td>
<td>final</td>
<td>bəth /bət/</td>
<td>path /pɑ:θ/</td>
<td>final</td>
</tr>
</tbody>
</table>

2. The use of (an) English symbol(s) to represent a Sinhala consonant sound in the word initial, medial and/or final positions where the particular symbol(s) produce(s) (a) different sound(s) in Sri Lankan English. Below are some examples:
IV. DISCUSSION

In his description of “frequent copying,” Johanson (2002) says that “elements which already exist in the basic code...undergo an increase or decrease in frequency of occurrence” in the basic code (p. 292). The first three main patterns of representing Sinhala vowel sounds and two main patterns of representing Sinhala consonant sounds can be considered what Johanson (2002, 2008) refers to as “frequent copies” (p. 292, p. 74) for the frequent spelling patterns in the model code English are copied onto the basic code Sinhala. The first and second sub-patterns under the fourth pattern of representing Sinhala vowel sounds seem to reflect the feature of “Netspeak” which Crystal (2004) talks about – the use of “repeated letters (aaaaahhhhh, hiiiiii, ooops, soooo) for emphasis” (p. 34). As shown in Crystal’s (2004) example, this feature occurs in the use of English on the Internet. Thus, they can also be considered “frequent copies.”

The copied English symbols are however, new to the basic code. Therefore, Johanson’s (2002, 2008) conception of “frequent copies” needs to be expanded. Accordingly, frequent copies can be referred to as elements which already exist in the basic code and/or are new to the basic code which may undergo an increase and/or decrease in the frequency of occurrence. In this respect, elements which already exist in the basic code may undergo either an increase or decrease in the frequency of occurrence while the elements that are new to the basic code may undergo an increase in the frequency of occurrence.

The research participants’ intense and regular use of the copied elements can be considered what Johanson calls “habitualization” (2002, p. 298, 2008, p. 65). As “habitualized copies”, the patterns of representing Sinhala vowel and consonant sounds through English symbols identified in the study tend to “occur frequently, regularly and normally” among the research participants’ texting (Johanson, 2002, p. 298). This frequent and regular use of the copies reflects a certain degree of “acceptance” of the copies by the bilingual research participants who seem to constitute a “speech community” (Johanson, 2002 p. 299, 2008 p. 65). However, it is necessary to observe the occurrence of the copies in the basic code over a longer period of time to identify them as “conventionalized code copies” (Johanson, 2002, p. 299, 2008, p. 65), for “the process of conventionalization is a continuum of changes in sociolinguistic status with gliding transitions between degrees of acceptability for individuals and for speech communities” (Johanson, 2002, p. 300). Thus, the study which focuses on messages sent by informants over a period of 10 months can argue that the copies seem to appear only as “part of a more general bilingual norm” (Johanson, 2002, p. 299) and seem to be undergoing a process of conventionalization.

In relation to Johanson’s (2002) claim that “language birth” may be the final result of conventionalization as the “high copying variety becomes the specific code of a whole speech community, including monolinguals” (p. 299), this study suggests that the linguistic phenomenon of using English letters to represent Sinhala vowel and consonant sounds in online text messages has potential to give birth to a new variety of Sinhala in which the orthographic system is a result of the contact between Sinhala and English orthography.

The patterns identified also challenge Gunasékara’s (1962) claim that “no system of transliteration of a phonetic language by an unphonetic one, like the English can ever be perfect” (p. 29). Just as there is a “method” in the “madness” or “inconsistencies” of English spelling (Vallins, 1965, p. 12), there is a method in the madness or inconsistencies in representing Sinhala sounds using English letters although they reflect the complexity of the contact between a more phonetic language like Sinhala and a less phonetic language like English (Fernando n.d., p. 10). Some patterns used to represent Sinhala vowel and consonant sounds through English letters reflect how the Sinhala-English bilingual Internet user takes advantage of the “madness” or “inconsistencies” (Vallins, 1965, p. 12) of English spelling to suit his/her own agenda either by using an existing English spelling pattern to represent a different sound or by creating a new spelling pattern. Certain patterns involve a reduplication of an English symbol in the representation of a long Sinhala vowel sound. It is presumably because the particular English symbol is also used to represent the short vowel sound of the said Sinhala vowel sound. It seems to reflect the Sinhala-English bilingual Internet user’s attempt to maintain the orthographic difference between the short and long Sinhala vowel sounds apparent in the Sinhala spelling system. This complexity of the representations of Sinhala sounds through English letters shows that the “unintelligibility” or the “ambiguity” of the Sinhala messages represented using English letters which

<table>
<thead>
<tr>
<th>English symbol used</th>
<th>Position in Sinhala words</th>
<th>Sinhala words</th>
<th>English words in which the particular symbol(s) produce(s) the Sinhala sound represented</th>
<th>The sounds produced by the said symbol in English words</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>initial</td>
<td>tibba /티바/  මීබා</td>
<td>-</td>
<td>/t/ - /tin/</td>
</tr>
<tr>
<td></td>
<td>medial</td>
<td>atta /اةට/  අටට</td>
<td>-</td>
<td>/t/ - /tin/</td>
</tr>
<tr>
<td></td>
<td>final</td>
<td>behet /බේහෙට/  බේහෙට</td>
<td>-</td>
<td>/t/ - /tin/</td>
</tr>
</tbody>
</table>

**Table 7. Spoken Sinhala dental unvoiced stop ‘ත’ /t̪/**
Widyalankara (2015, p. 1) refers to could be subjective for it could be a result of a particular user’s unfamiliarity with the patterns used by another user.

These complex patterns of representing Sinhala vowel and consonant sounds through English characters are also useful to help develop more user-friendly online Sinhala-English transliteration software than the present ones which are unable to identify certain patterns of representing Sinhala vowel and consonant sounds through English characters (i.e. ‘Google Input Tools’) identified in the study or have their own transliteration schemes including capitalized English letters and punctuation marks which the users have to adopt when typing online (i.e. ‘SinGlish Transliterated’ and ‘SinGlish (Phonetic) Transliterated’).

ACKNOWLEDGMENT
My sincere gratitude is extended to my research supervisor for her continuous guidance throughout the process of writing my undergraduate dissertation on which this paper is based, with a few modifications. I am also grateful to all my lecturers, colleagues and research participants. This research would not have been possible, if not for their support. Most of all, I am sincerely thankful to all my loved ones who always wish my happiness and success.

NOTES
1 Mendis (2006) notes that a “text message” is the popular name for SMS (Short Message Service) which is “a type of digital discourse that has come into being with the developments in mobile telephone technology” (p. 125). This study is on ‘online’ text messages which belong to “digital internet genres” (Mendis, 2006, p. 125).
2 Mobile messaging applications “allow you to send and receive pictures or text messages without paying for SMS” (Holt, Bossler & Seigfried, 2015, p. 341).
3 In the simplest definition, “language contact is the use of more than one language in the same place at the same time” (Thomason, 2001, p. 1).
4 Speech sounds are certain acoustic effects voluntarily produced by the organs of speech; they require that the speech-organs shall be placed in certain definite positions or moved in certain definite ways” (Jones, 1969, p. 1). Rajapaksha says that spoken Sinhala has thirteen vowel sounds (1997, p. 12). Unlike Disanayaka, Rajapaksha does not include the mid central long vowel sound /a/ in his classification of vowels but the mid central short vowel sound /a/ which he refers to as “schwa”. For a detailed illustration of the distribution of spoken Sinhala vowel and consonant sounds, see Disanayaka, 1991, pp. 65-124.
5 “Letters provide a means of symbolizing sounds” (Jones, 1956, p. 11). The word ‘letters’ will be used interchangeably with the words ‘symbols’ and ‘characters’ in the study.
6 Due to space constraints, the paper presents only a few examples of spoken Sinhala vowel and consonant sounds (for more examples, see Malalasekera, pp.19-69). Most examples of English words included in the tables are taken from Vallins (1965). For more examples, see Vallins, 1965, pp. 48-51. Note that Vallins (1965) uses IPA phonetic symbols while this study uses Fernando’s (1985) SLE phonetic symbols for it represents the Sri Lankan English pronunciation of the words. The phonetic symbols used by Weerasinghe et al. (2004-2007) are used to represent the Sinhala vowel and consonant sounds in the tables.

REFERENCES


